Long Run and Short Effects in Static Panel Models

Peter Egger and Michael Pfaffermayr*

January 28, 2002

Abstract

For short and fat panels the Mundlak model can be viewed as an approximation of a general dynamic autoregressive distributed lag model. We give an exact interpretation of short run and long effects and provide simulations to assess the quality of the approximation of the long run and short run effects by the parameters of the Mundlak Model.

Key words: Random Effects Models, Mundlak Model, Panel Econometrics.

JEL classification:

*Affiliation of both authors: Department of Economic Theory, Economic Poliy and Economic History, University of Innsbruck, Universitätsstr. 15, A-6020 Innsbruck, Austria and Austrian Institute of Economic Research, PO-Box 91, A-1103-Vienna, Austria, E-mail: Peter.Egger@uibk.ac.at and Michael.Pfaffermayr@uibk.ac.at.

1 Introduction[1]

Since the seminal work of Kuh (1959) and Houthakker (1965) it is argued that in static panel models the long run effects are mainly captured by between estimates, while the within estimates represent short-run effects.

Baltagi and Griffin (1984) investigate the case of underspecified lag dynamics in a finite distributed lag error components model. They conclude that "the OLS estimator provides a robust estimator of the long run ... elasticity under alternative degrees of misspecification, variance components, and time series observations. In contrast, the within estimator offers a good estimator of the short run effects but can severely underestimate the long run response." (ibid, p. 643)

Pirotte (1999) assumes a general dynamic error components model as the data generating process and investigates whether the within and between estimates approximate the short and long run effects properly. From his Monte Carlo simulations he concludes that "the probability limit of the Between estimator of the static model converges, in all cases, to long run effects" and that "long run effects are obtained directly from the static relation without the need of a dynamic model" (p. 155).

In many applications, data comprise a panel with a large cross-section dimension, but only a few observations over time. In such short and fat panels, it is impossible to estimate complex dynamic models and it is important to investigate, *whether* and *under which circumstances* static panel models provide a reasonable approximation of the short run and the long

[1] We are grateful to Sylvia Kaufmann and Andrea Weber for helpful comments.

run effects. The present paper follows Pirotte (1999) in assuming a dynamic error components model as the data generating process (more precisely an autoregressive distributed lag model ADL(1,1)). In contrast to previous research, we demonstrate that disregarding the dynamic process (the lagged endogenous variable) results in an approximation error and in autocorrelated residuals. Noteworthy, the former does not vanish in large cross-sections, which has been suggested in Pirotte (1999). We derive the probability limit of the approximation error of both the short run and the long run parameter estimate. Moreover, we assess the determinants of this approximation bias and its small sample properties in a Monte Carlo experiment. Specifically, we demonstrate that both the short run and the long run parameter are underestimated and this approximation bias is the larger, the slower the adjustment of the dependent variable (i.e. the higher the omitted parameter of the lagged dependent variable) and the shorter the time dimension of the panel.

The next section discusses the within and between estimates on the basis of the Mundlak (1978) model, which seems to be the natural specification in this case, and evaluates the corresponding approximation biases. Section 3 reports the results of the Monte Carlo simulations. The last section summarizes the main findings.

2 The basic model and the approximation bias

Mundlak's (1978) model can be interpreted as an approximation to an ADL(1,1) model (or ADL-models in general), providing both long run and short run parameter estimates. Without loss of generality, we confine our analysis to one exogenous variable and assume that the data generating process is given by the following dynamic error components model:

$$y_{it} = \gamma y_{i,t-1} + x_{it}\beta + x_{it-1}\pi + \mu + \alpha_i + \varepsilon_{it}. \tag{1}$$

x_{it} is the explaining variable, β is the corresponding parameter capturing the short run impact, while $\boldsymbol{\pi}$ represents the lagged impact of x_{it}. μ is the constant, the random individual effects are denoted by $\alpha_i \sim N(0, \sigma_\alpha^2)$ and the remainder error is $\varepsilon_{it} \sim N(0, \sigma_\epsilon^2)$. We assume that x_{it} is doubly exogenous (see Cornwell et al., 1992), i.e. $E[x_{it}\alpha_i] = 0$ and $E[x_{it}\varepsilon_{it}] = 0$. Furthermore, we only look at stationary data where both the left hand side variable y_{it} and the explaining right hand side variable x_{it} are $I(0)$. Hence, we do not consider either a unit root in x_{it} or y_{it} or a cointegrating relationship between them. In vector form, the model reads:

$$\mathbf{y} = \mathbf{y}_{-1}\gamma + \mathbf{x}\beta + \mathbf{x}_{-1}\boldsymbol{\pi} + \mu\boldsymbol{\iota}_{NT} + \mathbf{Z}_{\boldsymbol{\alpha}}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \tag{2}$$

where $\mathbf{y}$ is $(NT \times 1)$, $\mathbf{x}$ is $(NT \times 1)$, $\boldsymbol{\alpha}$ is a $(N \times 1)$ vector of random effects. $\mathbf{P} = \mathbf{I}_N \otimes \overline{\mathbf{J}}_T/T$ with $\overline{\mathbf{J}}_T$ is a $(T \times T)$ matrix of ones, $\mathbf{Z}_{\boldsymbol{\alpha}}$ is defined as $\mathbf{I}_N \otimes \boldsymbol{\iota}_T$. We define the lag polynomial $\frac{\Theta(L)}{\Gamma(L)} = \frac{\beta+\pi L}{1-\gamma L} = \sum_{\tau=0}^{\infty} \theta_\tau L^\tau$ with L denoting the lag operator and

$$\theta_0 = \beta,\ \theta_1 = \gamma\beta + \pi,\ \theta_2 = \gamma\theta_1,\ \theta_3 = \gamma\theta_2.... \tag{3}$$

This MA-representation is used to reformulate the basic model as in Pirotte (1999). In addition, we introduce the implied approximation error of the static error components model.

$$
\begin{aligned}
y_{it} &= \frac{\Theta(L)}{\Gamma(L)} x_{it} + \frac{\mu + \alpha_i + \varepsilon_{it}}{\Gamma(L)} \\
&= x_{it}\beta + \frac{1}{1-\gamma}\overline{x}_i(\gamma\beta + \pi) + \sum_{j=0}^{\infty} \gamma^j (x_{it-j-1} - \overline{x}_i)(\gamma\beta + \pi) \\
&\quad + \frac{\mu}{1-\gamma} + \frac{\alpha_i}{1-\gamma} + u_{it} \\
&= x_{it}\beta + \overline{x}_i\widetilde{\varphi} + \sum_{j=0}^{\infty} \gamma^j (x_{it-j-1} - \overline{x}_i)(\gamma\beta + \pi) + \widetilde{\mu} + \widetilde{\alpha}_i + u_{it}
\end{aligned} \tag{4}
$$

where $\widetilde{\varphi} = \frac{1}{1-\gamma}(\gamma\beta + \pi)$, $\widetilde{\mu} = \frac{\mu}{1-\gamma}$, $\widetilde{\alpha}_i = \frac{\alpha_i}{1-\gamma}$. The transformation with the lag polynomial $\Gamma(L)$ furthermore implies that the error term is now AR(1) and given by $u_{it} = \gamma u_{it-1} + \varepsilon_{it}$ (see Greene, 1993) with $\sigma_u^2 = \frac{\sigma_\varepsilon^2}{1-\gamma^2}$. We assume an infinite history of the explaining variable, so that the approximation error becomes $\sum_{j=0}^{\infty} \gamma^j (x_{it-j-1} - \overline{x}_i)(\gamma\beta + \pi)$. It can immediately be seen that this term vanishes, if $\gamma = 0$ and the underlying model reduces to that analyzed in Baltagi and Griffin (1984). Therefore, in the absence of a lagged dependent variable the Mundlak model is a perfect representation of a model with lagged exogenous variables, and the underspecified lag dynamics is fully compensated by the inclusion of the group mean as a control.

In contrast to Baltagi and Griffin (1984), Pirotte (1999), and others we do not consider a dynamic data generating process for the right hand side variable x_{it}, rather in both the theoretical analysis and the simulation exercise we assume that the x_{it} is IID, namely $x_{it} = \mu\zeta_i + \eta_{it}$ with $\mu \in R$, $\zeta_i \sim N\left(0, \sigma_\zeta^2\right)$, $\eta_{it} \sim N\left(0, \sigma_\eta^2\right)$, $E[\zeta_i\eta_{it}] = 0$.

Applying the within transformation $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ to the Mundlak approximation, one gets

$$\begin{aligned}
\mathbf{Qy} &= \mathbf{Qx}\beta + \mathbf{QPx}\widetilde{\varphi} + \sum_{j=0}^{\infty} \gamma^{j} \mathbf{Q}\left(\mathbf{x}_{-j-1} - \mathbf{Px}\right)(\gamma\beta + \pi) + \mathbf{QZ}_{\alpha}\widetilde{\boldsymbol{\alpha}} + \mathbf{Qu} \\
&= \mathbf{Qx}\beta + \sum_{j=0}^{\infty} \gamma^{j} \mathbf{Qx}_{-j-1}(\gamma\beta + \pi) + \mathbf{Qu}, \qquad (5)
\end{aligned}$$

where $\mathbf{x}_{-j-1}$ is $(NT \times 1)$ and includes the elements $(x_{i1-j-1}, x_{i2-j-1}, ..., x_{iT-j-1})'$. This approximation produces an omitted variable bias, which for fixed T and an infinite history of x_{it} is given by

$$p \lim_{N\to\infty} \left[\widehat{\beta}_{Within} - \beta\right] = p \lim_{N\to\infty} \left[\sum_{j=0}^{\infty} \gamma^{j} (\mathbf{x}'\mathbf{Qx})^{-1} \mathbf{x}'\mathbf{Qx}_{-j-1}(\gamma\beta + \pi)\right], \qquad (6)$$

since $p\lim_{N\to\infty}\left[\mathbf{x}'\mathbf{Qu}\right] = 0$, because x_{it} is exogenous and uncorrelated with the error term. For $j < T - t$, it follows that

$$\begin{aligned}
& p \lim_{N\to\infty} [\left(x_{it} - \frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau}\right)\left(x_{it-j-1} - \frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau-j-1}\right)] = \qquad (7) \\
& p \lim_{N\to\infty} [x_{it}x_{it-j-1} - x_{it}\frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau-j-1} - x_{it-j-1}\frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau} \\
& + \frac{1}{T^2}\sum_{\tau=1}^{T}\sum_{\tau'=1}^{T} x_{i\tau}x_{i\tau'-j-1}] = \\
& \mu^2\sigma_{\zeta}^2 - 2\mu^2\sigma_{\zeta}^2 - \frac{2}{T}\sigma_{\eta}^2 + \mu^2\sigma_{\zeta}^2 + \frac{T-j-1}{T^2}\sigma_{\eta}^2 = -\frac{T+j+1}{T^2}\sigma_{\eta}^2,
\end{aligned}$$

while it is zero otherwise $(j \geq T - t)$. Furthermore,

$$\begin{aligned}
p \lim_{N\to\infty} [\left(x_{it} - \frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau}\right)\left(x_{it} - \frac{1}{T}\sum_{\tau=1}^{T} x_{i\tau}\right)] &= \qquad (8) \\
\mu^2\sigma_{\zeta}^2 + \sigma_{\eta}^2 - 2\mu^2\sigma_{\zeta}^2 - \frac{2}{T}\sigma_{\eta}^2 + \mu^2\sigma_{\zeta}^2 + \frac{1}{T}\sigma_{\eta}^2 &= \frac{T-1}{T}\sigma_{\eta}^2,
\end{aligned}$$

Combining these two results, the probability limit of the approximation bias of the short run parameter (after within transformation) amounts to

$$p\lim_{N\to\infty}\left[\sum_{j=0}^{T-1}\gamma^{j}(\mathbf{x}^{'}\mathbf{Q}\mathbf{x})^{-1}\mathbf{x}^{'}\mathbf{Q}\mathbf{x}_{-j-1}\right](\gamma\beta+\pi)= \tag{9}$$
$$\sum_{j=0}^{T-1}\gamma^{j}p\lim_{N\to\infty}\left((\mathbf{x}^{'}\mathbf{Q}\mathbf{x})^{-1}\right)\cdot p\lim_{N\to\infty}\left(\mathbf{x}^{'}\mathbf{Q}\mathbf{x}_{-j-1}\right)(\gamma\beta+\pi)=$$
$$-\sum_{j=0}^{T-1}\gamma^{j}\left(\frac{T+j+1}{T\left(T-1\right)}\right)(\gamma\beta+\pi)<0.$$

For large N and fixed T, the within estimate tends to underestimate the true short run impact, especially, in short panels (small T) and with a high persistence parameter γ. However, it can immediately be seen that the bias vanishes, if T approaches infinity and as long as $\gamma<1$.[2]

The long run impact of the explaining variable can be calculated from the between transformation, since

$$E\left(y\right)\big|_{x_{it}=x_{it-1}=\overline{x}}=\frac{\beta+\pi}{1-\gamma}\overline{x} \tag{10}$$

and

$$\mathbf{Py}=\mathbf{Px}\beta+\mathbf{Px}\widetilde{\varphi}+\sum_{j=0}^{\infty}\gamma^{j}\mathbf{P}\left(\mathbf{x}_{-j-1}-\mathbf{Px}\right)(\gamma\beta+\pi)+\mathbf{PZ}_{\boldsymbol{\alpha}}\widetilde{\boldsymbol{\alpha}}+\mathbf{Pu}. \tag{11}$$

The corresponding approximation bias of the long run impact is

$$p\lim_{N\to\infty}\left[\widehat{\beta}+\widehat{\widetilde{\varphi}}-\beta-\widetilde{\varphi}\right]= \tag{12}$$
$$p\lim_{N\to\infty}\left[\sum_{j=0}^{\infty}\gamma^{j}\left[(\mathbf{x}^{'}\mathbf{Px})^{-1}\mathbf{x}^{'}\mathbf{Px}_{-j-1}-\mathbf{I}\right](\gamma\beta+\pi)\right],$$

[2] $\lim_{T\to\infty}\sum_{j=0}^{T-1}\gamma^{j}\left(\frac{T+j+1}{T(T-1)}\right)(\gamma\beta+\pi) < (\gamma\beta+\pi)\lim_{T\to\infty}\left[\left(\frac{2T+1}{T(T-1)}\right)\sum_{j=0}^{T-1}\gamma^{j}\right] = (\gamma\beta+\pi)\lim_{T\to\infty}\left(\frac{2T+1}{T(T-1)}\right)\lim_{T\to\infty}\sum_{j=0}^{T-1}\gamma^{j}=0$, since $\lim_{T\to\infty}\left(\frac{2T+1}{T(T-1)}\right)=0$.

since in addition to $p\lim_{N\to\infty}\left[\mathbf{x}^{'}\mathbf{Qu}\right] = 0$, $p\lim_{N\to\infty}\left[\mathbf{x}^{'}\mathbf{PZ}_{\boldsymbol{\alpha}}\widetilde{\boldsymbol{\alpha}}\right] = 0$ holds by assumption. For $j < T-1$, one obtains

$$p\lim_{N\to\infty}\left[\left(\frac{1}{T}\sum_{\tau=1}^{T}x_{i\tau}\right)\left(\frac{1}{T}\sum_{\tau=1}^{T}x_{i\tau-j-1}\right)\right] = \mu^2\sigma_\zeta^2 + \frac{T-j-1}{T^2}\sigma_\eta^2,$$

while for $j \geq T-1$ we have

$$p\lim_{N\to\infty}\left[\left(\frac{1}{T}\sum_{\tau=1}^{T}x_{i\tau}\right)\left(\frac{1}{T}\sum_{\tau=1}^{T}x_{i\tau-j-1}\right)\right] = \mu^2\sigma_\zeta^2. \tag{13}$$

Combining terms yields

$$p\lim_{N\to\infty}\left[\sum_{j=0}^{\infty}\gamma^j\left[(\mathbf{x}^{'}\mathbf{Px})^{-1}\mathbf{x}^{'}\mathbf{Px}_{-j-1} - \mathbf{I}\right](\gamma\beta+\pi)\right] = \tag{14}$$

$$\left[-\sum_{j=0}^{T-2}\gamma^j\left(\frac{(j+1)\sigma_\eta^2}{T(T\mu^2\sigma_\zeta^2+\sigma_\eta^2)}\right) - \sum_{j=T-1}^{\infty}\gamma^j\left(\frac{\sigma_\eta^2}{T\mu^2\sigma_\zeta^2+\sigma_\eta^2}\right)\right](\gamma\beta+\pi) < 0.$$

It is evident that in absolute terms the approximation bias of the between estimate shrinks, if T grows large, if the persistence parameter γ becomes smaller or if the between variation in the right hand side variable (σ_ζ^2) is high or becomes more and more important (μ increases). This bias of the long run estimate likewise tends to zero as T approaches infinity for $\gamma < 1$.[3]

The Mundlak model integrates both approaches in an error components framework. Transforming the data according to Fuller and Battese (1973) by $\sigma_\varepsilon\Omega^{-1/2} = \mathbf{Q} + \frac{\sigma_\varepsilon}{\sigma_1}\mathbf{P}$, $\sigma_1 = \sqrt{T\sigma_\alpha^2+\sigma_\varepsilon^2}$ gives the well known Mundlak (1978)

[3]It suffices to show that $\lim_{T\to\infty}\sum_{j=0}^{T-2}\gamma^j\left(\frac{(j+1)\sigma_\eta^2}{T(T\mu^2\sigma_\zeta^2+\sigma_\eta^2)}\right) = 0$. An upper bound is given by $\lim_{T\to\infty}\sum_{j=0}^{T-2}\gamma^j\left(\frac{(T+1)\sigma_\eta^2}{T(T\mu^2\sigma_\zeta^2+\sigma_\eta^2)}\right) = \lim_{T\to\infty}\left(\frac{(T+1)\sigma_\eta^2}{T(T\mu^2\sigma_\zeta^2+\sigma_\eta^2)}\right)\lim_{T\to\infty}\sum_{j=0}^{T-2}\gamma^j = 0$.

result, namely that one gets exactly the within and the between estimate for the short and the (additional) long run parameter, and the above stated results on the corresponding approximation biases apply:

$$\sigma_\varepsilon \Omega^{-1/2}\mathbf{y} = \left(\frac{\sigma_\varepsilon}{\sigma_1}\mathbf{P} + \mathbf{Q}\right)\mathbf{x}\beta + \left(\frac{\sigma_\varepsilon}{\sigma_1}\mathbf{P} + \mathbf{Q}\right)\mathbf{P}\mathbf{x}\widetilde{\varphi} + \tag{15}$$

$$\sum_{j=0}^{\infty}\gamma^j \left(\frac{\sigma_\varepsilon}{\sigma_1}\mathbf{P} + \mathbf{Q}\right)(\mathbf{x}_{-j-1} - \mathbf{P}\mathbf{x})(\gamma\beta + \pi) + \sigma_\varepsilon\Omega^{-1/2}(\mathbf{Z}_{\boldsymbol{\alpha}}\widetilde{\boldsymbol{\alpha}} + \mathbf{u})$$

$$= \mathbf{Q}\mathbf{x}\beta + \frac{\sigma_\varepsilon}{\sigma_1}\mathbf{P}\mathbf{x}(\beta + \widetilde{\varphi}) + \sum_{j=0}^{\infty}\gamma^j\mathbf{Q}\mathbf{x}_{-j-1}(\gamma\beta + \pi) +$$

$$\frac{\sigma_\varepsilon}{\sigma_1}\sum_{j=0}^{\infty}\gamma^j(\mathbf{P}\mathbf{x}_{-j-1} - \mathbf{P}\mathbf{x})(\gamma\beta + \pi) + \sigma_\varepsilon\Omega^{-1/2}(\mathbf{Z}_{\boldsymbol{\alpha}}\widetilde{\boldsymbol{\alpha}} + \mathbf{u}).$$

Ignoring the approximation bias, using the orthogonality between $\mathbf{P}$ and $\mathbf{Q}$ and the partitioned inverse, we have

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}}_{Mundlak} \\ \widetilde{\boldsymbol{\varphi}}_{Mundlak} \end{pmatrix} = \begin{pmatrix} (\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}\mathbf{y} \\ (\mathbf{X}'\mathbf{P}\mathbf{X})^{-1}\mathbf{X}'\mathbf{P}y - (\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}'\mathbf{Q}\mathbf{y} \end{pmatrix} \tag{16}$$

and

$$p\lim_{N\to\infty}\left[\begin{pmatrix} \widehat{\beta}_{Mundlak} \\ \widehat{\beta}_{Mundlak} + \widehat{\widetilde{\varphi}}_{Mundlak} \end{pmatrix} - \begin{pmatrix} \beta \\ \beta + \widetilde{\varphi} \end{pmatrix}\right] = \tag{17}$$

$$p\lim_{N\to\infty}\begin{pmatrix} \sum_{j=0}^{\infty}\gamma^j(\mathbf{x}'\mathbf{Q}\mathbf{x})^{-\mathbf{1}}\mathbf{x}'\mathbf{Q}\mathbf{x}_{-j-1}(\gamma\beta + \pi) \\ \sum_{j=0}^{\infty}\gamma^j(\mathbf{x}'\mathbf{P}\mathbf{x})^{-\mathbf{1}}(\mathbf{x}'\mathbf{P}\mathbf{x}_{-j-1} - \mathbf{x}'\mathbf{P}\mathbf{x})(\gamma\beta + \pi) \end{pmatrix}.$$

Summing up, we can view β as approximating the pure short run effect, while $\beta + \widetilde{\varphi}$ approximates the long run impact. Similar to others, the conclusion then is that in short panels the Mundlak-Model (the within and the between estimate) forms a natural approximation of a dynamic data generating process (ADL-model). This approximation induces AR(1) errors, and estimating an AR(1) model in the spirit of Baltagi and Li (1991), besides improving efficiency, provides some valuable information on the extent of the approximation bias, since it gives a direct (though in small samples not necessarily consistent) estimate of γ.

Another observation is that the familiar Hausman (1978) test may lead to a rejection of the static random effects model even if the "true" model is a dynamic random effects one (i.e. $\pi = 0$, $\gamma > 0$). The reason is that the approximation induces correlation between the explaining variable and the individual random effect due to the omission of the long run effect. Hence, the rejection of a test of $\widetilde{\boldsymbol{\varphi}}_{Mundlak} = 0$ can have three reasons. (i) $\widehat{\widetilde{\boldsymbol{\varphi}}} \neq 0$ although $\gamma = 0$ and $\pi = 0$ (no dynamics), which is the traditional Hausman (1978) and Mundlak (1978) result of testing $E[x_i \alpha] = 0$. $E[x_i \alpha] = 0$, but the true $\widetilde{\boldsymbol{\varphi}} \neq 0$ because of omitting a relevant long run impact with (ii) $\pi \neq 0$ and/or $\gamma > 0$ or (iii) $\pi = 0$ and $\gamma > 0$. To our knowledge, the latter two reasons have not been considered before.

3 Simulation Results

Our Monte Carlo simulation set-up compares four model parameters to assess the sensitivity of four different estimators: the error components model (Amemiya, 1971, type), which ignores the long run effects, the Mundlak model, which obtains both the within estimates and the between estimates, the error components AR(1) model and the Mundlak AR(1) model using the autocorrelation correction proposed by Baltagi and Li (1991).[4] We assume

[4] In the Mundlak AR(1) model, the group means do not represent pseudo averages over time. This results in slight deviations of the estimated between error component from its

that $\beta = \pi = 1$, $\sigma_\varepsilon = \sigma_\eta = \sigma_\zeta = 1$ and vary the other underlying parameters in the following form:

- Strength of the autoregressive process, i.e the coefficient of the lagged dependent variable: $\gamma = 0.2,\ 0.8$.
- The relative importance of the between variance component in the data generating process of y_{it} : $\theta = 1 - \sqrt{\frac{\sigma_\varepsilon}{T\sigma_\alpha + \sigma_\varepsilon}} = 0.5,\ 0.9$, so that σ_α is implicitly defined.
- Time dimension: $T = 5,\ 10,\ 20$.
- Cross-section versus time variation in the data generating process of x_{it}: $x_{it} = \mu\zeta_i + \eta_{it}$, $\mu = 1, 2$.

We set $N = 100$ and replicate each experiment 10000 times. Since we are estimating four specifications in each case, we end up with 24 simulation experiments. Table 1 contains the results for $\mu = 1$, Table 2 for $\mu = 2$. The tables provide information on both the true parameters, the average estimated parameter in each model (Av.) as well as bias, standard deviation (Std) and root mean squared error (Rmse) of both the estimated short run (β) and long run parameters ($\beta + \widetilde{\varphi}$). Since our focus is on short panels, we only briefly discuss the effects of an increase in T.

*** Insert Tables 1 and 2 about here ***

In line with Pirotte (1999), we find that the sum of the between estimator approximates the long run effect quite well for all analyzed time dimensions,

random effects counterpart.

especially, if γ is low. This seems to hold independently of the relative importance of the cross-section variation as represented by θ and μ. However, if γ is large, the approximation bias turns out to be substantial (an underestimation of up to 25% on average) much larger than the variance of the parameter estimate. Noteworthy, the autocorrelation correction tends to reduce the standard errors only marginally, but not the RMSE, which slightly increases in most cases.

In contrast, the within estimate approximates the short run impact only weakly, particularly so if γ is large. The approximation gets better with rising T and it seems independent of θ or μ. Here, the autocorrelation correction does not improve the precision of the estimates, rather it aggravates the downward bias, especially for small T and high γ. Hence, in short panels, the within estimate can be recommended to obtain short run estimates (β) if the autoregressive process is not too strong (γ is low).

For the simple random effects model, which ignores the additional long run effect, two sources of the bias emerge. First, there is the downward approximation bias. Second, omitting the long run impact induces an upward bias (for $\pi > 0$), since the random effects estimate is a weighted average of the between and the within estimate. This leads to the paradox result that in some cases the random effects model outperforms the within estimates. Especially, if γ is high, one may be better off with the random effects model. If the cross-section variation is important, the random effects AR(1) performs best (high θ and/or high μ), but the within estimator is superior if θ is low. However, for $\gamma = 0.2$ and $\theta = 0.9$ (even more so if $\mu = 2$) the random effects model overestimates the short run effect substantially. For example, for $\gamma = 0.2$, $\theta = 0.9$, $\mu = 2$ and $T = 10$ the average bias is more than 500 %.

4 Conclusions

The estimation of short run and long run effects in static panels with a short time series dimension as an approximation of a dynamic model depends crucially on the parameter of the lagged dependent variable of the underlying autoregressive distributed lag (ADL) model. The probability limits for large N and fixed T imply that both the short run and the long run effect are downward biased. The bias turns out to be the stronger the higher the parameter of the lagged dependent variable, the shorter the time dimension and, as far as the long run impact is concerned, the more important the between variation in the explaining variable.

The simulation exercise confirms these results for small samples and shows that for a sufficiently fast adjustment of the dependent variable (low γ) both the within and between estimates produce reasonable approximations of the short and long run effects even in panels with a short time series dimension. In all constellations of the Monte Carlo simulation set-up, the between estimates as proxy of the long run effects perform considerably better than the within estimates as proxy of the short run effects. The correction for the autocorrelation in the remainder error term, which is induced by this approximation, does not improve the precision of the estimates and cannot be recommended in all cases. Especially with short time series, AR(1) estimation even aggravates the approximation bias of the within estimate.

For practitioners, our analysis suggests that one can use the Mundlak model to approximate short run and long run effects, when inference in a dynamic model is not feasible. To assess the quality of this approximation, it seems important to check the importance of the induced autocorrelation of the errors (γ). The latter can maybe not precisely estimated, but it nevertheless allows to assess the strength of the underlying dynamic process. Long run and, especially, short run effects are always strongly underestimated, if γ is high. The size of the between variation seems less important. We would not recommend the standard random effects model as an approximation. Although it outperforms the Mundlak model in terms of the short run estimate in some cases, it substantially overstimates the short run effect, especially, if the within variation is relatively important. There are two biases simultaneously at work (the omitted long run effects bias *and* the approximation bias), and it is difficult to predict their net effect. Despite the induced autocorrelation of the errors, parameter inference from AR(1) corrected models seems not a good strategy in short panels, especially, if the induced autocorrelation (γ) is strong.

5 References:

Amemiya, T., 1971, The Estimation of the Variances in a Variance-Components Model, *International Economic Review* 12, pp. 1-13.

Baltagi, B.H., Griffin, J.H., 1984, Short and Long Run Effects in Pooled Models, *International Economic Review* 25(3), 631-645,.

Baltagi, B.H., Li, Q., A Transformation that Will Circumvent the Problem of Autocorrelation in an Error Components Model, *Journal of Econometrics* 48, pp. 385-393.

Cornwell, Ch., Schmidt, P., Wyhowski, D. J., 1992, Simultaneous Equations and Panel Data, *Journal of Econometrics* 51, pp. 151-181.

Fuller, W.A., Battese, G.E., 1973, Transformations for Estimation of Linear Models with Nested Error Structure, *Journal of the American Statistical Association* 68, pp. 626-632.

Greene, W., 1993, Econometric Analysis, Macmillan, New York, 2nd edition.

Hausman, J.A., 1978, Specification Tests in Econometrics, *Econometrica* 46, pp. 1251-1271.

Houthakker, H.S, 1965, New Evidence on Demand Elasticities, *Econometrica* 33, pp. 277-288.

Kuh, E., 1959, The Validity of Cross-sectionally Estimated Behaviour Equations in Time Series Applications, *Econometrica* 27, pp. 197-214.

Mundlak, Y., 1978, On the Pooling of Time Series and Cross Section Data, *Econometrica* 46(1), pp. 69-85.

Pirotte, A., 1999, Convergence of the Static Estimation Toward Long Run Effects of the Dynamic Panel Data Models, *Economics Letters* 63, pp. 151-158.

Table 1 - Monte Carlo Results I: x_{it}=ζ_i+η_{it}

		β				$\beta+\tilde{\varphi}$							True parameters					
		Av.	Bias	Std	Rmse	Av.	Bias	Std	Rmse	σ_u	$\sigma_{\tilde{\alpha}}$	ρ	T	β	$\beta+\tilde{\varphi}$	γ	θ	$\sigma_{\tilde{\alpha}}$
1	REM	1.61	0.61	0.12	0.63					1.12	1.01		5	1.0	2.5	0.2	0.5	0.97
	Mundlak	0.72	0.28	0.11	0.30	2.48	0.02	0.11	0.11	1.12	1.01							
	REM AR(1)	1.66	0.66	0.13	0.67					1.36	0.96	0.17						
	Mundlak AR(1)	0.63	0.37	0.11	0.38	2.48	0.02	0.11	0.16	1.36	0.94	0.17						
2	REM	1.39	0.39	0.29	0.49					1.53	3.64		5	1.0	10.0	0.8	0.5	3.87
	Mundlak	0.26	0.74	0.15	0.75	7.61	2.39	0.36	2.41	1.53	3.64							
	REM AR(1)	1.21	0.21	0.27	0.34					2.55	4.86	0.59						
	Mundlak AR(1)	0.14	0.86	0.09	0.87	7.57	2.43	0.36	2.54	2.55	3.44	0.59						
3	REM	0.79	0.21	0.11	0.24					1.12	5.55		5	1.0	2.5	0.2	0.9	5.56
	Mundlak	0.72	0.28	0.11	0.30	2.48	0.02	0.56	0.56	1.12	5.55							
	REM AR(1)	0.74	0.26	0.11	0.28					1.38	5.49	0.16						
	Mundlak AR(1)	0.63	0.37	0.11	0.38	2.48	0.02	0.56	0.57	1.38	5.48	0.16						
4	REM	0.36	0.64	0.20	0.67					2.04	18.37		5	1.0	10.0	0.8	0.9	22.25
	Mundlak	0.26	0.74	0.20	0.76	7.62	2.38	1.80	2.99	2.04	18.37							
	REM AR(1)	0.22	0.78	0.11	0.78					2.48	18.56	0.68						
	Mundlak AR(1)	0.13	0.87	0.11	0.88	7.57	2.43	1.79	3.10	2.48	18.04	0.68						
5	REM	1.66	0.66	0.10	0.67					1.16	0.70		10	1.0	2.5	0.2	0.5	0.68
	Mundlak	0.85	0.15	0.08	0.16	2.49	0.01	0.08	0.08	1.16	0.70							
	REM AR(1)	1.52	0.52	0.12	0.54					1.34	0.76	0.21						
	Mundlak AR(1)	0.68	0.32	0.07	0.33	2.49	0.01	0.08	0.20	1.34	0.64	0.21						
6	REM	1.82	0.82	0.31	0.88					1.90	2.76		10	1.0	10.0	0.8	0.5	2.74
	Mundlak	0.45	0.55	0.12	0.57	8.48	1.52	0.28	1.55	1.90	2.76							
	REM AR(1)	0.26	0.74	0.06	0.74					1.73	6.81	0.75						
	Mundlak AR(1)	0.13	0.87	0.06	0.88	8.32	1.68	0.28	1.87	1.73	2.38	0.75						
7	REM	0.92	0.08	0.08	0.11					1.16	3.93		10	1.0	2.5	0.2	0.9	3.93
	Mundlak	0.85	0.15	0.08	0.16	2.49	0.01	0.40	0.40	1.16	3.93							
	REM AR(1)	0.77	0.23	0.08	0.25					1.35	3.90	0.21						
	Mundlak AR(1)	0.68	0.32	0.08	0.33	2.49	0.01	0.40	0.44	1.35	3.88	0.21						
8	REM	0.54	0.46	0.14	0.48					2.23	13.91		10	1.0	10.0	0.8	0.9	15.73
	Mundlak	0.45	0.55	0.14	0.57	8.50	1.50	1.39	2.04	2.23	13.91							
	REM AR(1)	0.15	0.85	0.06	0.86					1.67	15.14	0.80						
	Mundlak (AR1)	0.12	0.88	0.06	0.88	8.32	1.68	1.36	2.30	1.67	13.45	0.80						
9	REM	1.68	0.68	0.09	0.69					1.17	0.49		20	1.0	2.5	0.2	0.5	0.48
	Mundlak	0.93	0.07	0.05	0.09	2.50	0.00	0.06	0.06	1.17	0.49							
	REM AR(1)	1.36	0.36	0.10	0.37					1.33	0.67	0.24						
	Mundlak AR(1)	0.70	0.30	0.05	0.31	2.50	0.00	0.06	0.24	1.33	0.44	0.24						
10	REM	2.15	1.15	0.31	1.19					2.11	2.05		20	1.0	10.0	0.8	0.5	1.94
	Mundlak	0.64	0.36	0.09	0.37	9.18	0.82	0.22	0.85	2.11	2.05							
	REM AR(1)	0.15	0.85	0.04	0.85					1.43	7.94	0.81						
	Mundlak AR(1)	0.12	0.88	0.04	0.88	8.91	1.09	0.21	1.37	1.43	1.57	0.81						
11	REM	0.98	0.02	0.05	0.06					1.17	2.78		20	1.0	2.5	0.2	0.9	2.78
	Mundlak	0.93	0.07	0.05	0.09	2.50	0.00	0.28	0.28	1.17	2.78							
	REM AR(1)	0.78	0.22	0.06	0.23					1.33	2.79	0.23						
	Mundlak AR(1)	0.70	0.30	0.05	0.31	2.50	0.00	0.28	0.36	1.33	2.74	0.23						
12	REM	0.73	0.27	0.10	0.29					2.27	10.42		20	1.0	10.0	0.8	0.9	11.12
	Mundlak	0.64	0.36	0.10	0.37	9.18	0.82	1.06	1.34	2.27	10.42							
	REM AR(1)	0.13	0.87	0.04	0.87					1.42	12.78	0.83						
	Mundlak AR(1)	0.11	0.89	0.04	0.89	8.89	1.11	1.03	1.72	1.42	9.97	0.83						

The true σ_α is transformed by 1/(1-γ) to get the variance components of the approximating model. The transformed σ_ϵ amounts to 1.02 and 1.67, respectively.

Table 2 - Monte Carlo Results II: x_{it}=$2\zeta_i$+η_{it}

		β				$\beta+\tilde{\varphi}$							True parameters					
		Av.	Bias	Std	Rmse	Av.	Bias	Std	Rmse	σ_u	$\sigma_{\tilde{\alpha}}$	ρ	T	β	$\beta+\tilde{\varphi}$	γ	θ	$\sigma_{\tilde{\alpha}}$
1	REM	2.14	1.14	0.08	1.14					1.12	1.01		5	1.0	2.5	0.2	0.5	0.97
	Mundlak	0.72	0.28	0.11	0.30	2.50	0.00	0.06	0.06	1.12	1.01							
	REM AR(1)	2.22	1.22	0.07	1.23					1.45	0.86	0.16						
	Mundlak AR(1)	0.63	0.37	0.11	0.38	2.50	0.00	0.06	0.11	1.45	0.94	0.16						
2	REM	4.31	3.31	0.55	3.36					2.03	3.60		5	1.0	10.0	0.8	0.5	3.87
	Mundlak	0.27	0.73	0.20	0.76	7.75	2.25	0.19	2.26	2.03	3.60							
	REM AR(1)	4.11	3.11	0.48	3.14					3.73	4.25	0.67						
	Mundlak AR(1)	0.13	0.87	0.10	0.88	7.70	2.30	0.19	2.40	3.73	3.21	0.67						
3	REM	0.97	0.03	0.11	0.12					1.12	5.56		5	1.0	2.5	0.2	0.9	5.56
	Mundlak	0.72	0.28	0.11	0.30	2.50	0.00	0.28	0.28	1.12	5.56							
	REM AR(1)	1.02	0.02	0.14	0.14					1.46	5.48	0.16						
	Mundlak AR(1)	0.63	0.37	0.11	0.38	2.50	0.00	0.28	0.30	1.46	5.49	0.16						
4	REM	0.77	0.23	0.28	0.36					2.43	18.34		5	1.0	10.0	0.8	0.9	22.25
	Mundlak	0.26	0.74	0.24	0.78	7.74	2.26	0.94	2.44	2.43	18.34							
	REM AR(1)	0.85	0.15	0.27	0.31					3.75	18.31	0.71						
	Mundlak AR(1)	0.12	0.88	0.12	0.88	7.69	2.31	0.93	2.58	3.75	17.96	0.71						
5	REM	2.15	1.15	0.07	1.15					1.16	0.70		10	1.0	2.5	0.2	0.5	0.68
	Mundlak	0.85	0.15	0.08	0.16	2.50	0.00	0.04	0.04	1.16	0.70							
	REM AR(1)	2.18	1.18	0.07	1.18					1.44	0.63	0.21						
	Mundlak AR(1)	0.68	0.32	0.08	0.33	2.50	0.00	0.04	0.18	1.44	0.64	0.21						
6	REM	5.13	4.13	0.56	4.17					2.51	2.71		10	1.0	10.0	0.8	0.5	2.74
	Mundlak	0.45	0.55	0.16	0.58	8.52	1.48	0.14	1.49	2.51	2.71							
	REM AR(1)	0.57	0.43	0.11	0.44					2.52	10.09	0.83						
	Mundlak AR(1)	0.11	0.89	0.06	0.89	8.32	1.68	0.14	1.83	2.52	1.83	0.83						
7	REM	1.07	0.07	0.08	0.11					1.16	3.93		10	1.0	2.5	0.2	0.9	3.93
	Mundlak	0.85	0.15	0.08	0.17	2.50	0.00	0.20	0.20	1.16	3.93							
	REM AR(1)	1.02	0.02	0.10	0.10					1.44	3.87	0.21						
	Mundlak AR(1)	0.68	0.32	0.08	0.33	2.50	0.00	0.20	0.27	1.44	3.88	0.21						
8	REM	0.99	0.01	0.23	0.23					2.76	13.90		10	1.0	10.0	0.8	0.9	15.73
	Mundlak	0.45	0.55	0.18	0.58	8.51	1.49	0.71	1.65	2.76	13.90							
	REM AR(1)	0.27	0.73	0.07	0.73					2.42	17.03	0.85						
	Mundlak (AR1)	0.11	0.89	0.06	0.89	8.30	1.70	0.70	1.97	2.42	13.33	0.85						
9	REM	2.16	1.16	0.06	1.16					1.17	0.49		20	1.0	2.5	0.2	0.5	0.48
	Mundlak	0.93	0.07	0.05	0.09	2.50	0.00	0.03	0.03	1.17	0.49							
	REM AR(1)	2.11	1.11	0.07	1.11					1.43	0.51	0.24						
	Mundlak AR(1)	0.70	0.30	0.05	0.31	2.50	0.00	0.03	0.23	1.43	0.44	0.24						
10	REM	5.51	4.51	0.56	4.54					2.66	2.02		20	1.0	10.0	0.8	0.5	1.94
	Mundlak	0.64	0.36	0.12	0.38	9.19	0.81	0.11	0.82	2.66	2.02							
	REM AR(1)	0.18	0.82	0.04	0.82					1.86	13.62	0.87						
	Mundlak AR(1)	0.11	0.89	0.04	0.89	8.84	1.16	0.10	1.36	1.86	0.78	0.87						
11	REM	1.13	0.13	0.06	0.15					1.17	2.78		20	1.0	2.5	0.2	0.9	2.78
	Mundlak	0.93	0.07	0.05	0.09	2.50	0.00	0.14	0.14	1.17	2.78							
	REM AR(1)	1.01	0.01	0.08	0.08					1.43	2.75	0.24						
	Mundlak AR(1)	0.70	0.30	0.05	0.31	2.50	0.00	0.14	0.27	1.43	2.74	0.24						
12	REM	1.14	0.14	0.18	0.23					2.79	10.42		20	1.0	10.0	0.8	0.9	11.12
	Mundlak	0.64	0.36	0.12	0.38	9.19	0.81	0.53	0.97	2.79	10.42							
	REM AR(1)	0.15	0.85	0.04	0.85					1.81	17.02	0.88						
	Mundlak AR(1)	0.11	0.89	0.04	0.89	8.83	1.17	0.51	1.45	1.81	9.82	0.88						

The true σ_α is transformed by 1/(1-γ) to get the variance components of the approximating model. The transformed σ_ϵ amounts to 1.02 and 1.67, respectively.